Reconciling qualitative and abstract (and scalable) modelling of biological regulatory networks

Loïc Paulevé

CNRS/LaBRI, Bordeaux, France

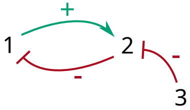
https://loicpauleve.name

joint work with
- T. Chatain, S. Haar, J. Kolčák (LSV/Inria, ENS Cachan, France)

Qualitative vs abstract modelling

information

Influence graph

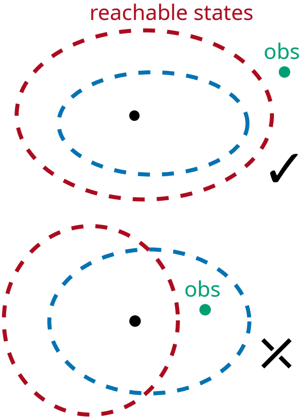


Boolean network
- logic of activity w.r.t. regulators
- *update mode* (sync, async, etc.)

Multilevel network
+ define activation thresholds

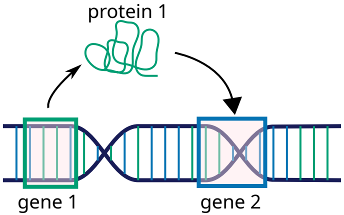
Quantitative model

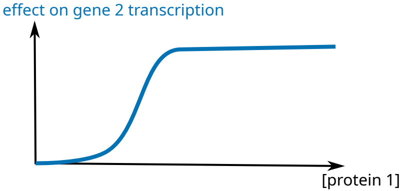
Consistency
analysis at Boolean level
transposable to multilevel?



Gene expression is not Boolean

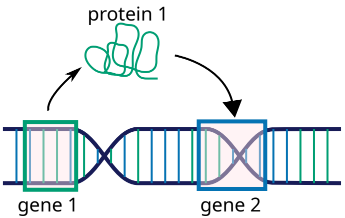
Qualitative modelling: Boolean vs multivalued networks

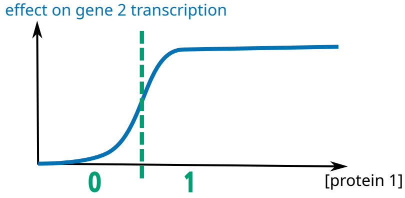




Gene expression is not Boolean

Qualitative modelling: Boolean vs multivalued networks





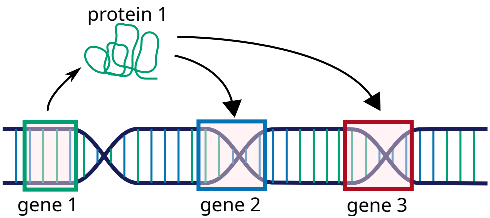
Boolean network

$$f : \mathbb{B}^n \to \mathbb{B}^n$$

$$f_2(x) \triangleq x_1$$

Gene expression is not Boolean

Qualitative modelling: Boolean vs multivalued networks

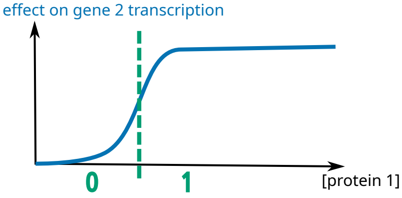


Boolean network

$$f : \mathbb{B}^n \to \mathbb{B}^n$$

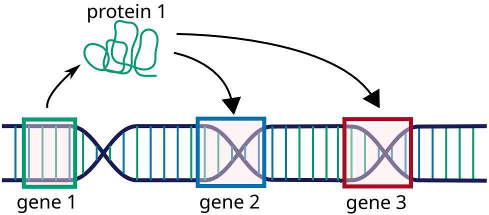
$$f_2(x) \triangleq x_1$$

$$f_3(x) \triangleq x_1$$



Gene expression is not Boolean

Qualitative modelling: Boolean vs multivalued networks

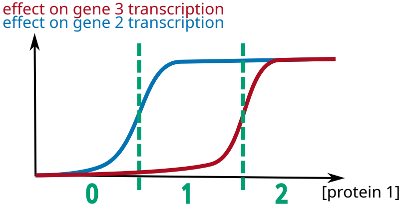


Boolean network

$$f : \mathbb{B}^n \to \mathbb{B}^n$$

$$f_2(x) \triangleq x_1$$

$$f_3(x) \triangleq x_1$$



Multivalued network

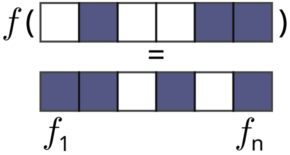
$$F_2(x) \triangleq (x_1 \geq 1)$$

$$F_3(x) \triangleq (x_1 \geq 2)$$

Introduction

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

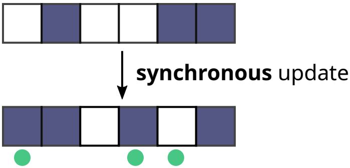
Configuration: $x \in \mathbb{B}^n$

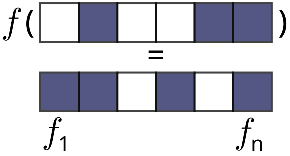


Introduction

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

Configuration: $x \in \mathbb{B}^n$

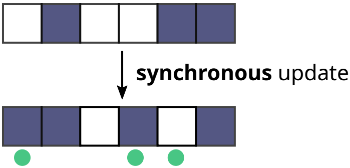


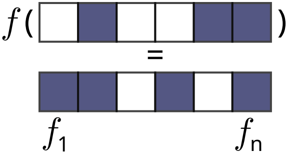


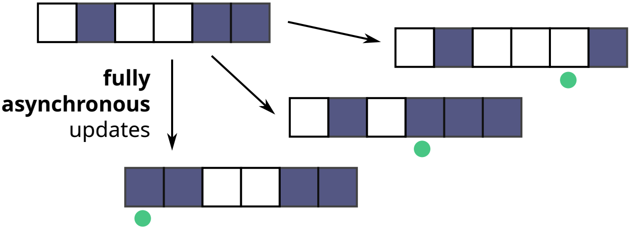
Introduction

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

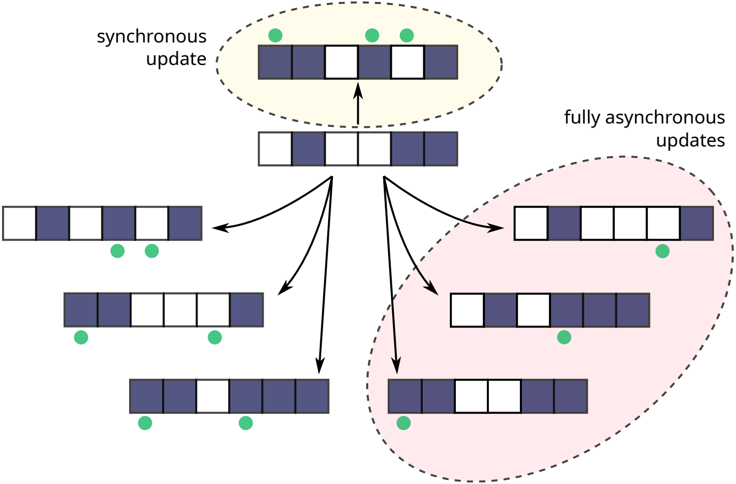
Configuration: $x \in \mathbb{B}^n$



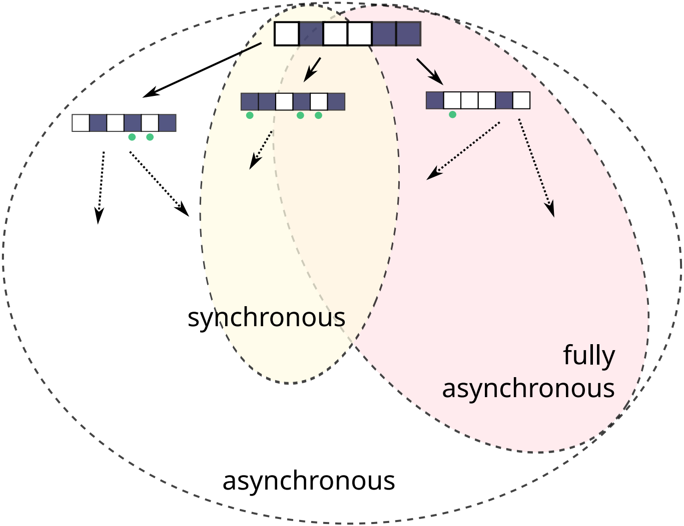




Asynchronous updates



Reachable configurations



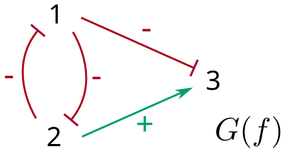
Asynchronous

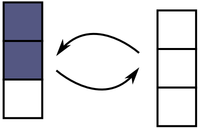
Boolean networks:

inconsistent abstraction

Motivating example
(embedded in many actual biological networks)

$$f_1(x) \triangleq \neg x_2$$
$$f_2(x) \triangleq \neg x_1$$
$$f_3(x) \triangleq \neg x_1 \wedge x_2$$



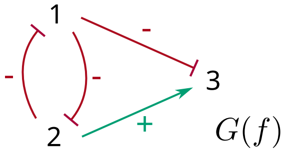


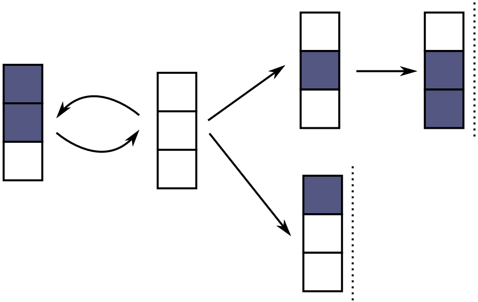
Motivating example
(embedded in many actual biological networks)

$$f_1(x) \triangleq \neg x_2$$

$$f_2(x) \triangleq \neg x_1$$

$$f_3(x) \triangleq \neg x_1 \wedge x_2$$





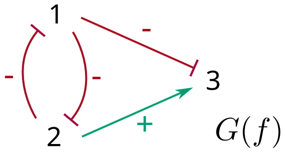
Motivating example

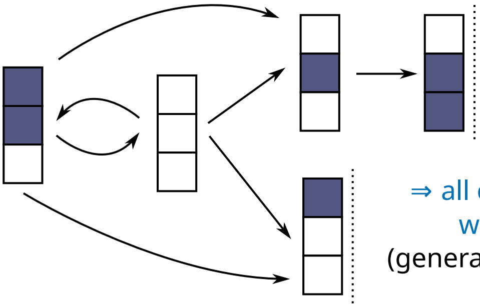
(embedded in many actual biological networks)

$$f_1(x) \triangleq \neg x_2$$

$$f_2(x) \triangleq \neg x_1$$

$$f_3(x) \triangleq \neg x_1 \wedge x_2$$

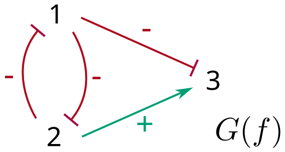




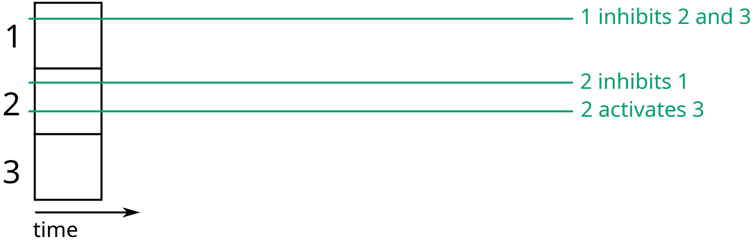
⇒ all configurations reachable
with any update mode
(generalized) asynchronous mode

Motivating example
(embedded in many actual biological networks)

$$F_1(x) \triangleq \neg(x_2 \geq 2)$$
$$F_2(x) \triangleq \neg x_1$$
$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$

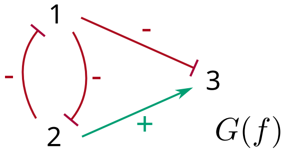


Compatible continuous/multilevel model:

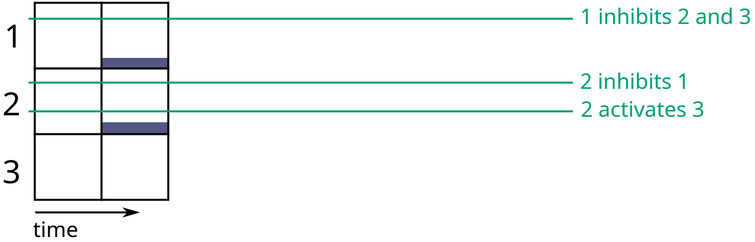


Motivating example
(embedded in many actual biological networks)

$$F_1(x) \triangleq \neg(x_2 \geq 2)$$
$$F_2(x) \triangleq \neg x_1$$
$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$

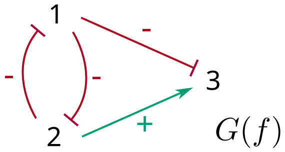


Compatible continuous/multilevel model:

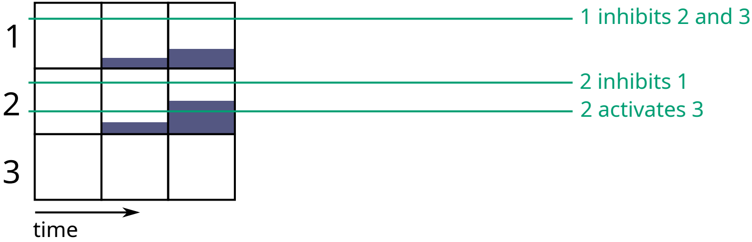


Motivating example
(embedded in many actual biological networks)

$$F_1(x) \triangleq \neg(x_2 \geq 2)$$
$$F_2(x) \triangleq \neg x_1$$
$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$



Compatible continuous/multilevel model:



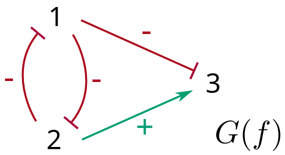
Motivating example

(embedded in many actual biological networks)

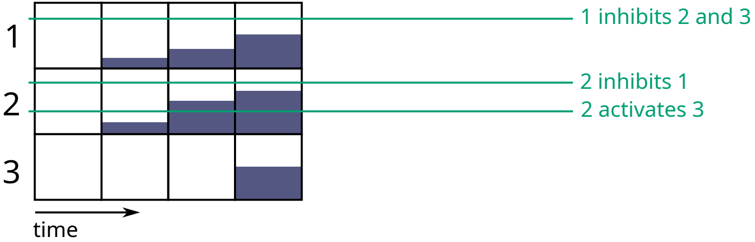
$$F_1(x) \triangleq \neg(x_2 \geq 2)$$

$$F_2(x) \triangleq \neg x_1$$

$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$

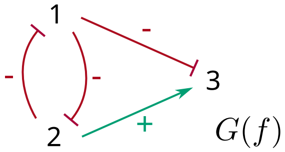


Compatible continuous/multilevel model:

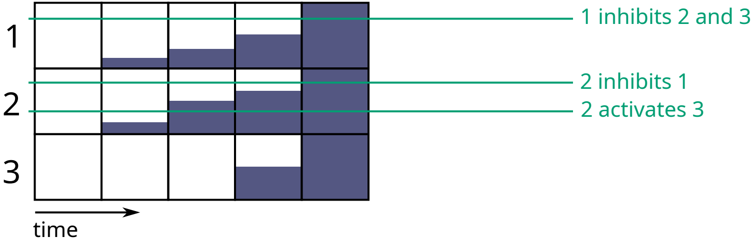


Motivating example
(embedded in many actual biological networks)

$$F_1(x) \triangleq \neg(x_2 \geq 2)$$
$$F_2(x) \triangleq \neg x_1$$
$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$



Compatible continuous/multilevel model:

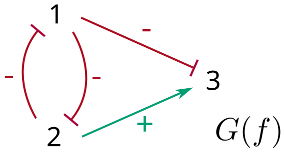


Motivating example
(embedded in many actual biological networks)

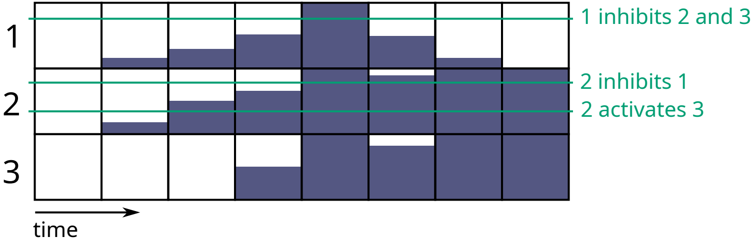
$$F_1(x) \triangleq \neg(x_2 \geq 2)$$

$$F_2(x) \triangleq \neg x_1$$

$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$

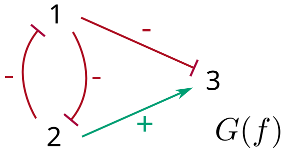


Compatible continuous/multilevel model:

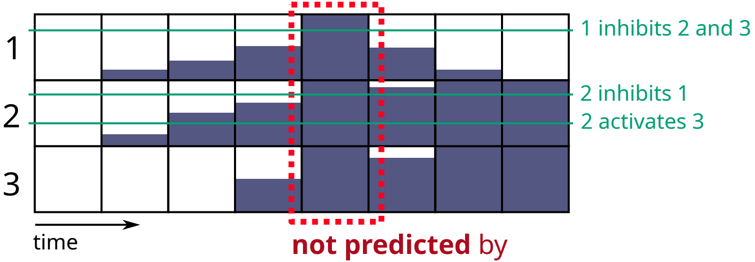


Motivating example
(embedded in many actual biological networks)

$$F_1(x) \triangleq \neg(x_2 \geq 2)$$
$$F_2(x) \triangleq \neg x_1$$
$$F_3(x) \triangleq \neg x_1 \wedge (x_2 \geq 1)$$



Compatible continuous/multilevel model:



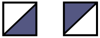
not predicted by
update modes in Boolean

Most Permissive
Boolean networks
enabling new behaviours

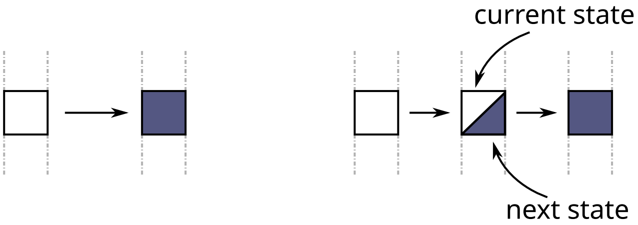
Most permissive semantics

- delay between firing and application of state change

  ⇒ allow interleaving other state changes

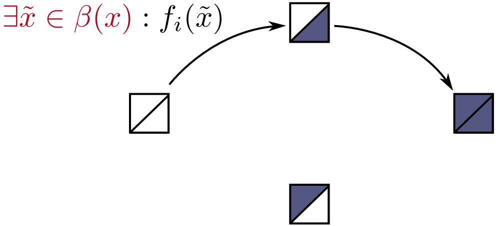
- in "intermediate" states 

  other components choose what they see



Most permissive semantics

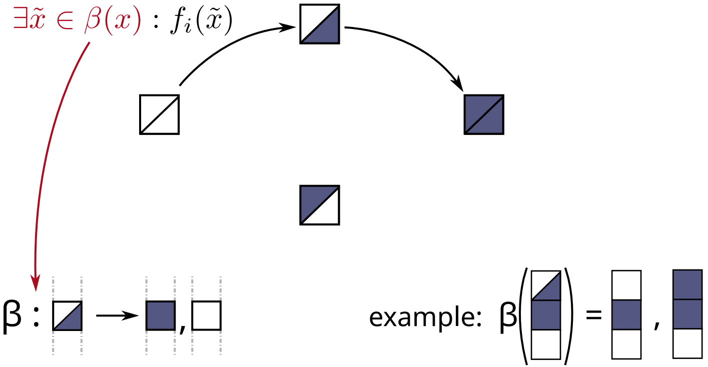
Automaton of component i:



+ fully-asynchronous interleaving

Most permissive semantics

Automaton of component i:

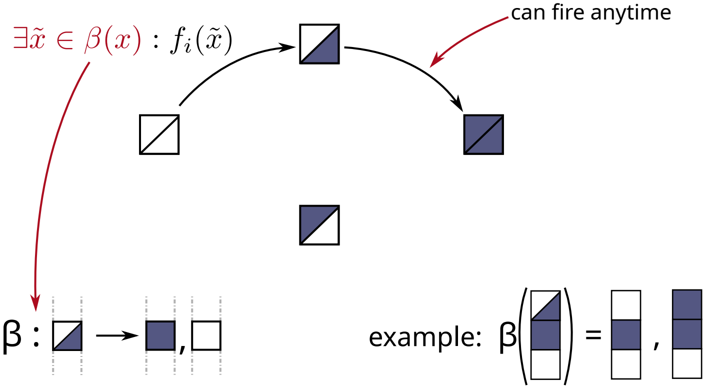


Choose value of "changing" components
(act as choosing an activation threshold)

+ fully-asynchronous interleaving

Most permissive semantics

Automaton of component i:

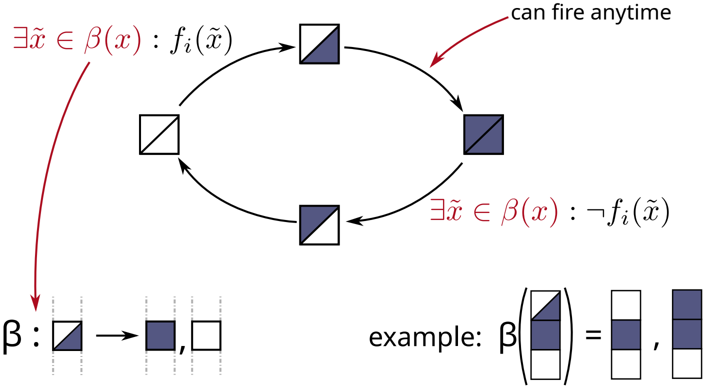


Choose value of "changing" components
(act as choosing an activation threshold)

+ fully-asynchronous interleaving

Most permissive semantics

Automaton of component i:

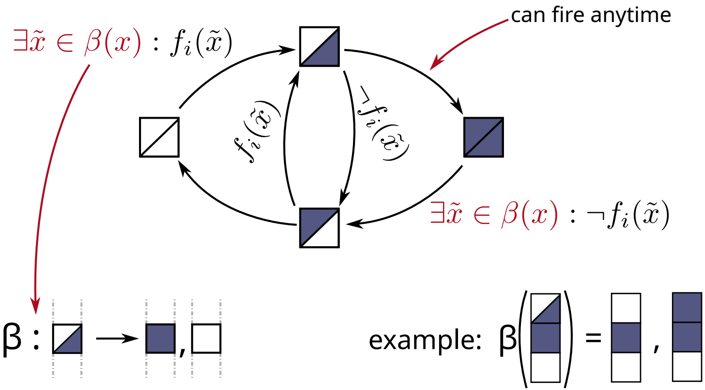


Choose value of "changing" components
(act as choosing an activation threshold)

+ fully-asynchronous interleaving

Most permissive semantics

Automaton of component i:



Choose value of "changing" components
(act as choosing an activation threshold)

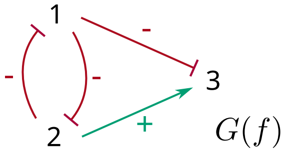
+ fully-asynchronous interleaving

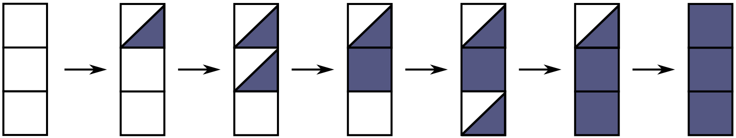
Application to motivating example

$$f_1(x) \triangleq \neg x_2$$

$$f_2(x) \triangleq \neg x_1$$

$$f_3(x) \triangleq \neg x_1 \wedge x_2$$





Properties of the most permissive semantics

Correct abstraction of multilevel/quantitative systems:

- includes all the transitions of every update mode
- quantitative refinements only remove behaviours

- Reachability can be decided in quadratic nb of transitions

⇒ y is reachable from x iff there is a path of length at most 3n

(PTIME with locally-monotonic networks, or encoded as BDDs/Petri nets/...;

NP-complete otherwise; instead of PSPACE-complete with update modes)

- Attractors are hypercubes (minimal trap spaces)

  ⇒ existence within a sub-space is NP-complete

  (instead of PSPACE-complete)

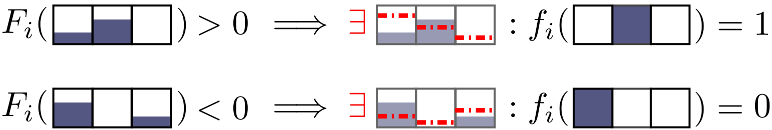
  ⇒ fixpoints are the same as with update modes

Refinements of Boolean Networks

A multivalued network

$$F : \mathbb{M}^n \to \{-1, 0, 1\}^n$$

is a refinement of a Boolean network *f* iff



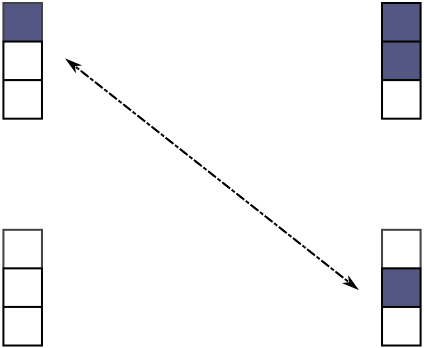
Most permissive semantics weakly simulates
any multivalued refinement with *any update mode*

(can be extended to ODEs)

Attractors of the most permissive semantics

Attractor: smallest set of configurations closed by transitions

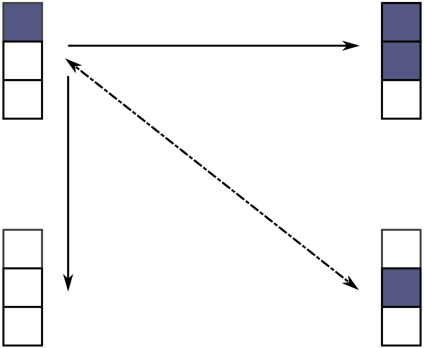
Attractors are hypercubes



Attractors of the most permissive semantics

Attractor: smallest set of configurations closed by transitions

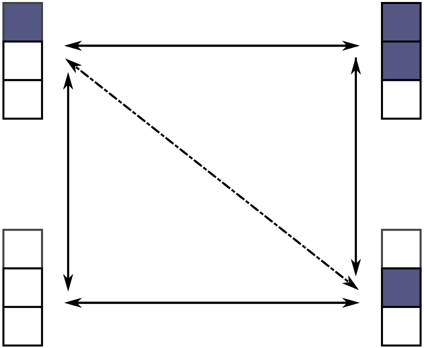
Attractors are hypercubes



Attractors of the most permissive semantics

Attractor: smallest set of configurations closed by transitions

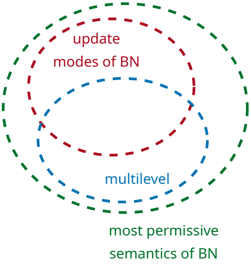
Attractors are hypercubes



Attractors of most permissive semantics = minimal trap spaces

Deciding if trap space is minimal = 2 reachability checking

Is most permissive semantics restrictive?



Minimality of abstraction

to any "most permissive" transition,

there is corresponding multilevel transition

Work in progress w/ "most permissive" paths:

non-minimal, but tricky counter-examples

- fixpoints are preserved (identical)
- trap spaces: known to be relevant for reasoning with attractors

  [Klarner et al in Nat. Comp. 2015] [Naldi in Front. Phys. 2018]

⇒ most permissive semantics seems still adequate to model differentiation processes !

Conclusion

Update modes of Boolean networks (sync, async, etc.):

• difficult to justify for biology (strong implications on dynamics)

• can miss important behaviours

⇒ lead to reject valid models of biological systems...

• have limited tractability (model-checking, ...)

Most permissive semantics:

• correct abstraction: guarantees that multilevel/quantitative refinements only remove behaviours

• simpler complexity: reachability PTIME, attractors NP

⇒ genome-scale tractability (BNs with >10,000 components)

Scalable Boolean analysis consistent with quantitative models